With the advent of post-genomic era and increasing use of computational techniques, great development has taken place in computational prediction in metabolic pathways which has numerous applications in systems biology and synthetic biology, for example, the design of tracer or knockout experiments (Blum and Kohlbacher, 2008). The primary method adopted by current tools for studying the metabolic pathways is to use the traditional algorithms of paths searching to compute the metabolic routes. Related works include Rahnuma (Mithani et al., 2005), MetaRoute (Blum and Kohlbacher, 2008), Pathway Hunter Tool (Rahman et al., 2005), FMM (Chou et al., 2009) and so on. Unlike previous methods, our approach takes into account the relative conservation of reactions that can be measured by the reaction’s frequency which is equals to the number of organisms that the reaction exists. Also, these tools either have no intermediates constraint or do not give a global and visual view of the search result. Furthermore, the user may want to design a route manually when there is no route to fulfill his needs in the search results; unfortunately, there are currently no tools provide such function, including the tools mentioned above.

Here, we described a tool called MRSD (Metabolic Route Search and Design) for searching and designing the metabolic routes based on the different organisms’ metabolic networks in KEGG. The search function enables users to perform route search in a weighted compound transforming graph in which the weight of the edge is a relevance function about the reaction’s frequency (Supplementary Material), in single organisms or groups of organisms with several constraints (such as using the intermediate metabolites or not). The new metabolic routes in the search results can be broadly used in synthetic biology. The design function can be employed when there is no metabolic route that satisfied the users’ need in the results computed by search function, or in another situation that the user only knows or only cares for the metabolic route from a specific source or end up with a specific product compound. The design function, which produces user’s interested routes in an interactive mode, can be used in biopharmaceuticals and other related fields.

2 METHODS

The MRSD functions based on the compound transforming graph which is a weighted digraph representing a metabolic network. The vertices in the graph represent the compounds, and the edges are the reactions connecting the substrates and products of the reactions. The head vertex acts as substrate in the reaction represented by the edge and the tail vertex acts as product. In this graph, one edge may represent various reactions because different reactions may have the same subset of reactants and products. In the graph of an individual organism, the weight of each edge is equal to the sum of frequencies of reactions that produce the corresponding product with the reactant. For a combined graph of various organisms, the weight of each edge is the sum frequencies of that in these organisms without regard to the duplicate of the reactions which may exists in more than one organism. The reason for this strategy is that we consider these reactions which exist in most of organisms have higher priorities than the others which only appear in few organisms.

MRSD contains two major functions: MRS (Metabolic Route Search) and MRD (Metabolic Route Design). On one hand, if you want to produce a known product from a known source, then the MRS can be employed. On the other hand, if only one metabolite is known and you want to find how to produce this metabolite or which metabolites can be produced by this one, then the MRD can be used in this situation. In addition, if the result routes of MRS do not fulfill your needs exactly, you can use MRD to design a route to introduce some minority modifies on the routes provided by MRS, which means that you can optimize the route by MRD according to the routes computed by MRS (Supplementary Material).

MRSD computes pathways between the source metabolite and the product metabolite by performing the k-shortest algorithm (Eppstein, 1998) on the weighted compound transform graph. MRS finds these shortest routes with maximum weight. Under these constraints, we convert the weighted graph to a probability weighted graph in which the edge with higher weight gets lower value in order to satisfied Eppstein’s algorithm (Supplementary Material). The result routes are a connected sequence of distinct reactions in which the product of one reaction acts as a substrate in the next reaction.
MRD designs a route in an interactive mode according to the assigned metabolite and the users’ choices. All the corresponding candidate metabolites are provided by MRD based on the user-preferred compound in each step. The computation of the candidates can be introduced to find the neighbour nodes of the current compound in the compound transfer graph.

3 APPLICATIONS

3.1 MRS

The main application of the MRS is exploring the appointed metabolic network for all routes transforming the source metabolite to the product metabolite, which can be extensively used in synthetic biology and other related fields.

There are several settings to be defined before starting a search: the source metabolite, product metabolite and the selected networks are required; and the remaining settings including whether the searched routes contain one metabolite or not are optional. The user can specify a list of intermediate metabolites separated by vertical line. The user can perform a route search on the metabolic network of one bacterium or on the combined network of a group of bacteria. Furthermore, it is possible for the user to search all routes in the super bacteria’ of the KEGG bacteria database. The ‘super bacteria’ is a virtual bacterium which is combined from all the bacteria in the KEGG bacteria database and contains all the reactions and compounds in this database. The user can also search the chemosynthesis route between two compounds represented by InChI string (Supplementary Material) by changing the interface type.

We perform a search on the metabolic network of the ‘super bacteria’ to find a route from glucose to 1-butanol (Atsumi et al., 2008). This is an interesting and useful route, which can facilitate strain improvement for specificity and productivity, as its production 1-butanol offers many advantages as a substitute for gasoline and diesel fuel. We perform a search on the metabolic network of the ‘super-bacteria’. (Fig. 1A) The designed route to produce succinate. Enzymatic reactions are drawn in orange.

4 CONCLUSIONS

The MRSD is a user-friendly and versatile tool that can be employed in many different applications including predicting biosynthetic or catabolic routes, biopharmaceuticals and other related fields through the separate or complementary use of MRS and MRD. MRS performs a route searching in the weighted compound transforming graph which take into account of the statistical property of the reactions, with several constraints considered to produce more convincing results. MRD is a new and useful tool to assist user in designing a route interactively. MRS and MRD can complement each other in helping researchers to continue their work conveniently.

ACKNOWLEDGEMENTS

The authors would like to thank Zhou Hong for advice about editing the article.

Funding: National Key Technologies R&D Program (2011CB910200).

Conflict of Interest: none declared.